Applications and Framework

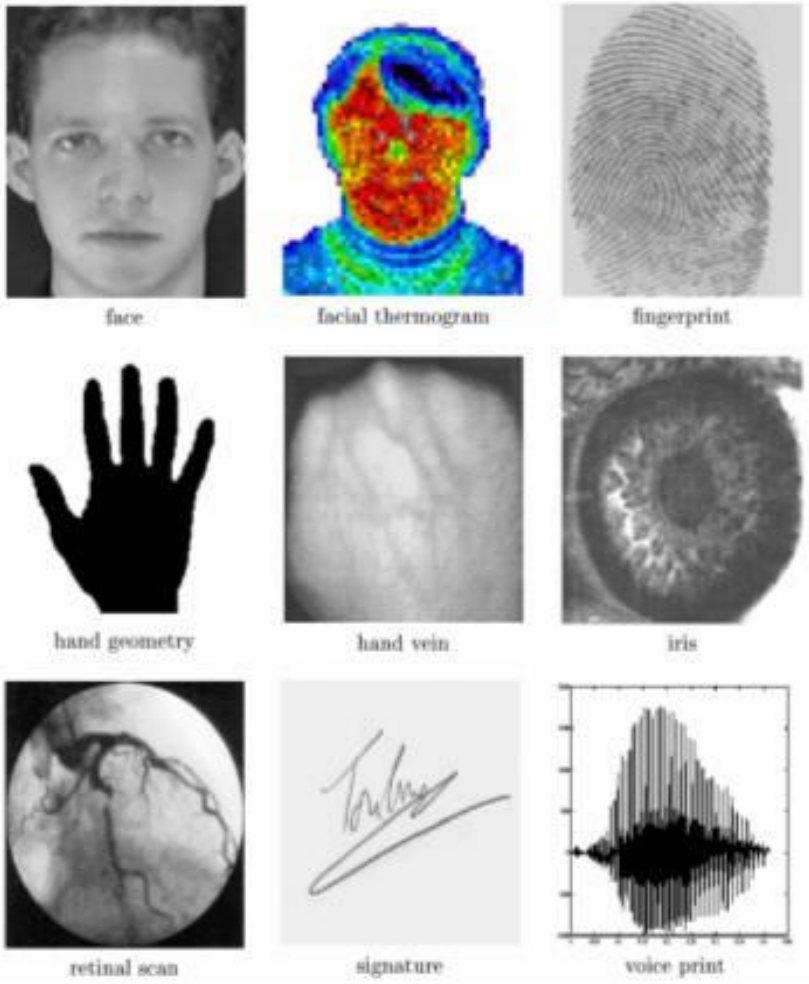
Dr. Fayyaz ul Amir Afsar Minhas
PIEAS Biomedical Informatics Research Lab
Department of Computer and Information Sciences
Pakistan Institute of Engineering & Applied Sciences
PO Nilore, Islamabad, Pakistan
http://faculty.pieas.edu.pk/fayyaz/

Applications of Machine Learning

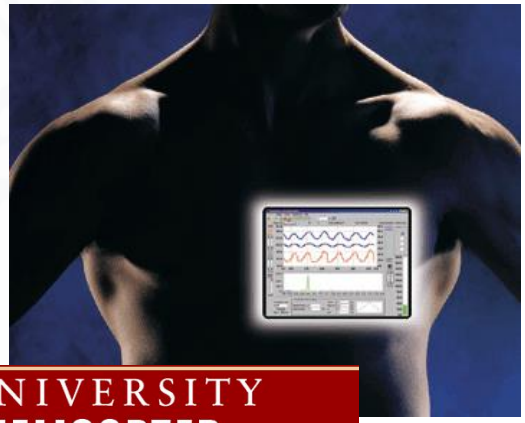
- An ability that I would like you to learn is to identify how to use machine learning in different domains.

- Machine learning can be applied in a wide array of real-world applications

Applications: Biometrics



Applications



STANFORD UNIVERSITY
AUTONOMOUS HELICOPTER

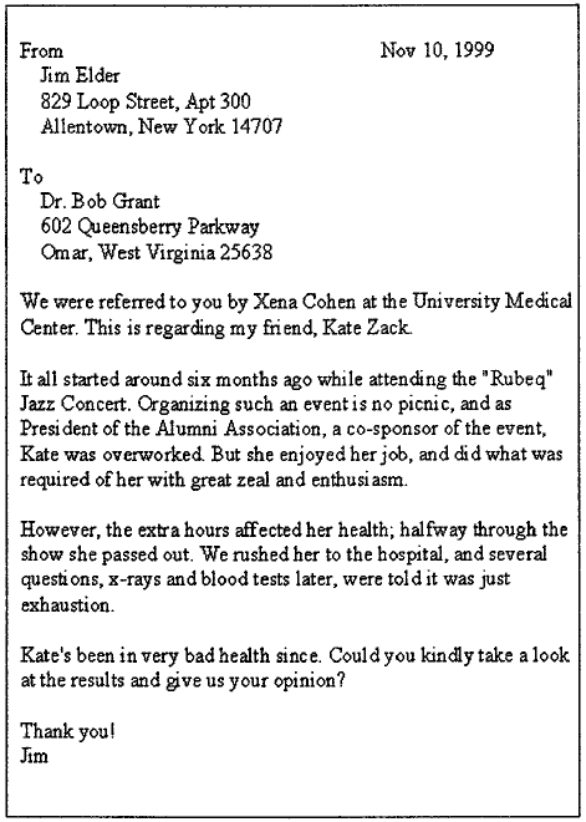
Overview

The goal of this project is to push the state-of-the-art in autonomous helicopter flight: extreme aerobatics under computer control.

Featured Videos

http://heli.stanford.edu/

Handwriting Recognition / OCR



From
Jim Elder
829 Loop Street, Apt 300
Allentown, New York 14707

Nov 10, 1999

To
Dr. Bob Grant
602 Queensberry Parkway
Omar, West Virginia 25638

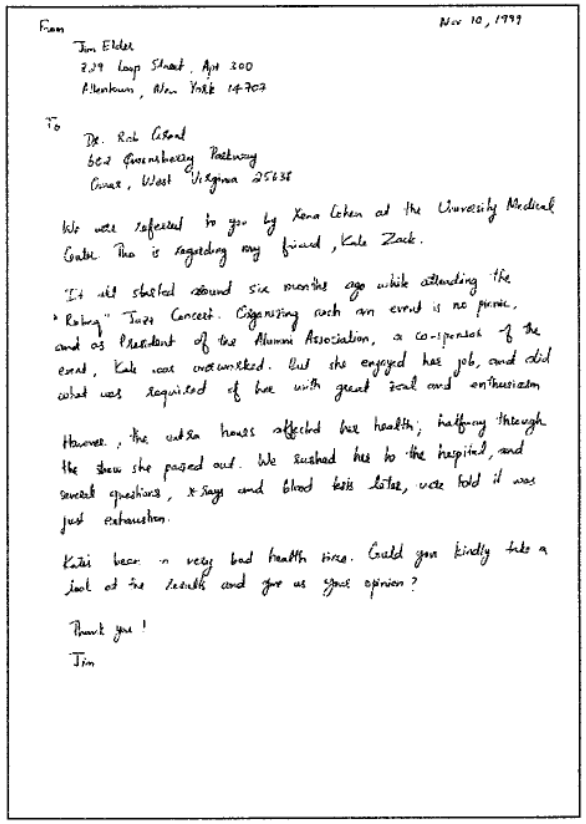
We were referred to you by Xena Cohen at the University Medical Center. This is regarding my friend, Kate Zack.

It all started around six months ago while attending the "Rubeq" Jazz Concert. Organizing such an event is no picnic, and as President of the Alumni Association, a co-sponsor of the event, Kate was overworked. But she enjoyed her job, and did what was required of her with great zeal and enthusiasm.

However, the extra hours affected her health; halfway through the show she passed out. We rushed her to the hospital, and several questions, x-rays and blood tests later, were told it was just exhaustion.

Kate's been in very bad health since. Could you kindly take a look at the results and give us your opinion?

Thank you!
Jim



From
Jim Elder
829 Loop Street, Apt 300
Allentown, New York 14707

Nov 10, 1999

To
Dr. Rob Grant
602 Queensberry Parkway
Omar, West Virginia 25638

We were referred to you by Xena Cohen at the University Medical Center. This is regarding my friend, Kate Zack.

It all started around six months ago while attending the "Rubeq" Jazz Concert. Organizing such an event is no picnic, and as President of the Alumni Association, a co-sponsor of the event, Kate was overworked. But she enjoyed her job, and did what was required of her with great zeal and enthusiasm.

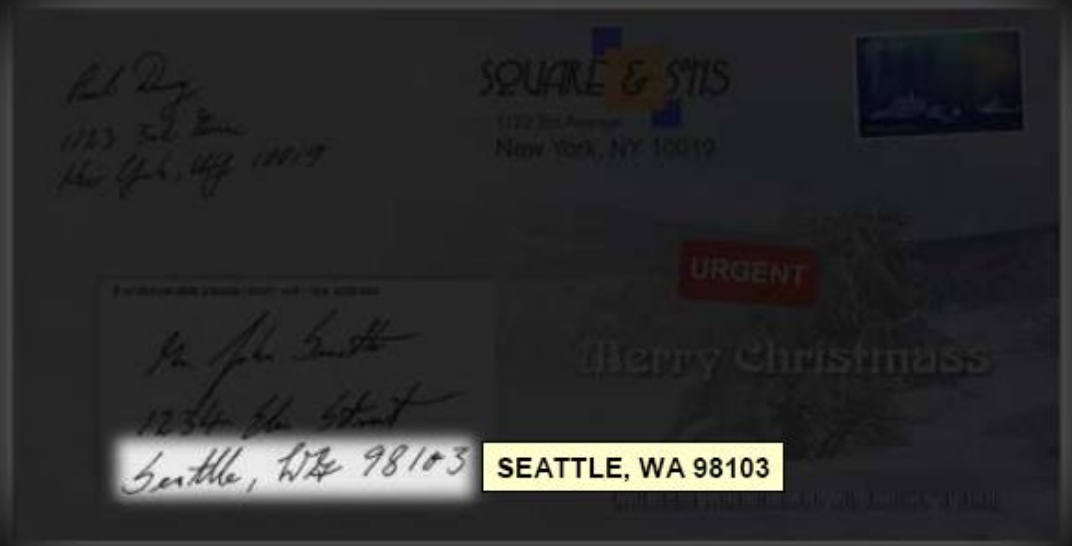
However, the extra hours affected her health; halfway through the show she passed out. We rushed her to the hospital, and several questions, x-rays and blood tests later, were told it was just exhaustion.

Kate's been in very bad health since. Could you kindly take a look at the results and give us your opinion?

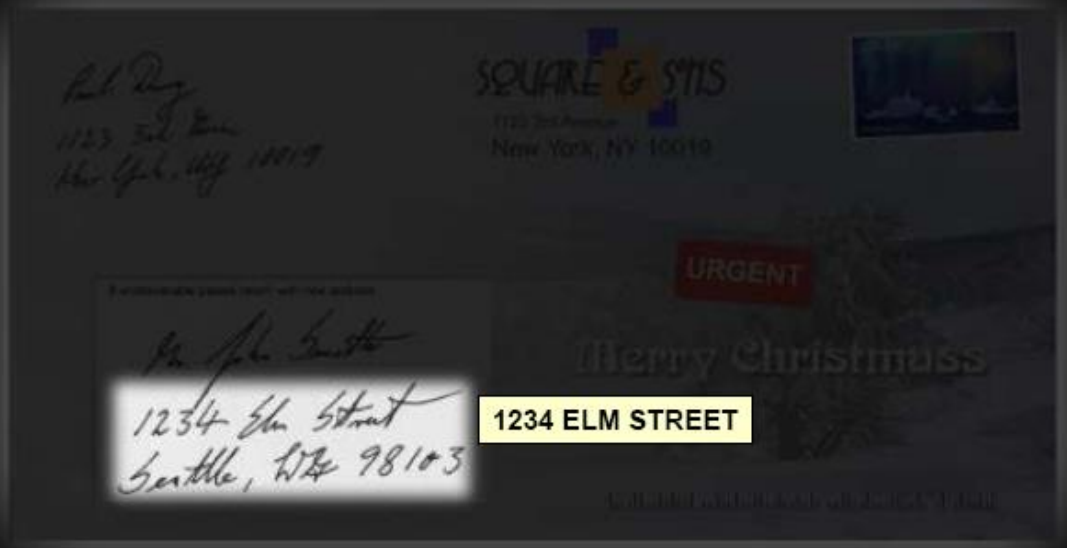
Thank you!
Jim

The Letter

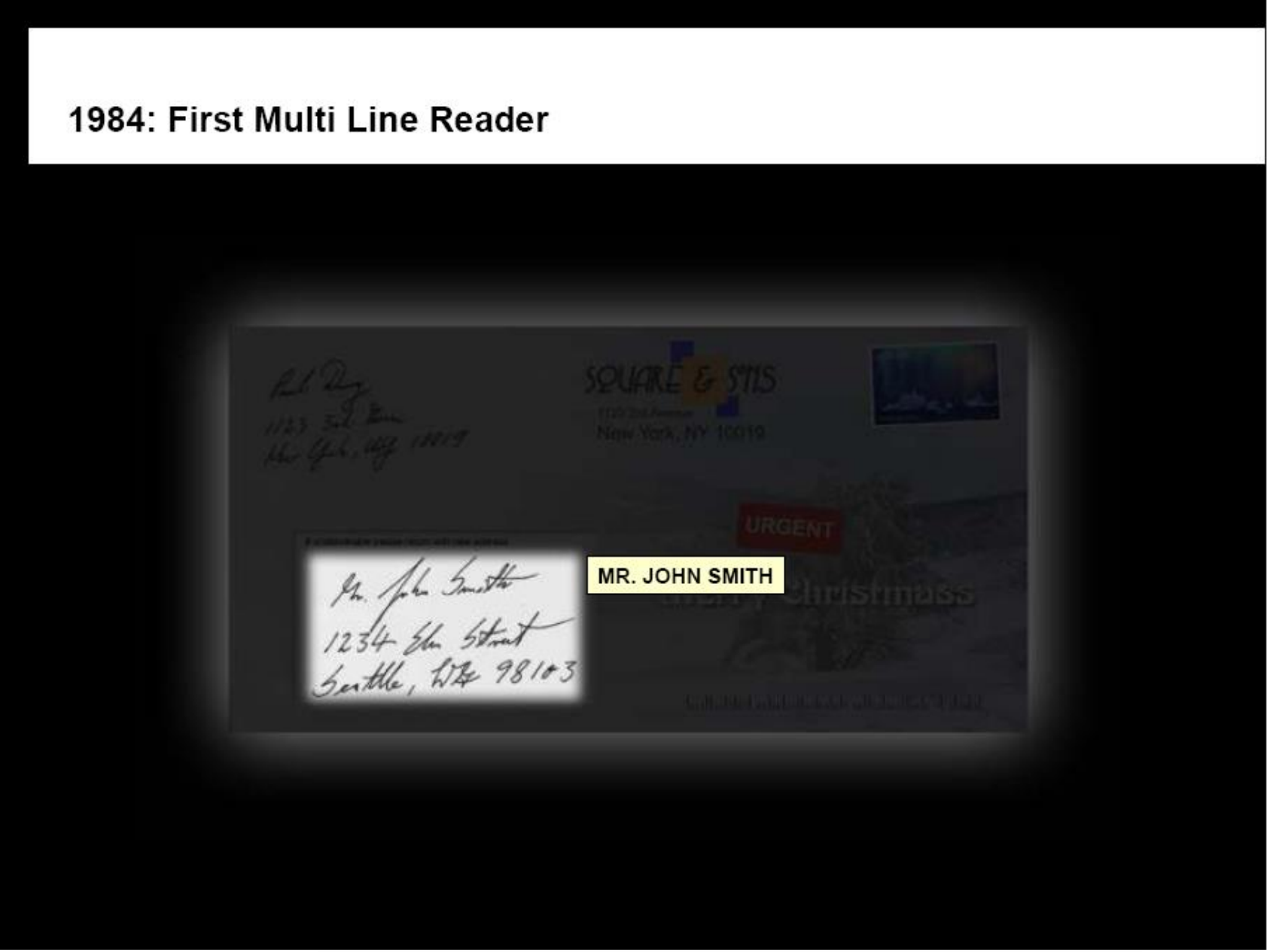
1978: First Postal Code Reader Worldwide



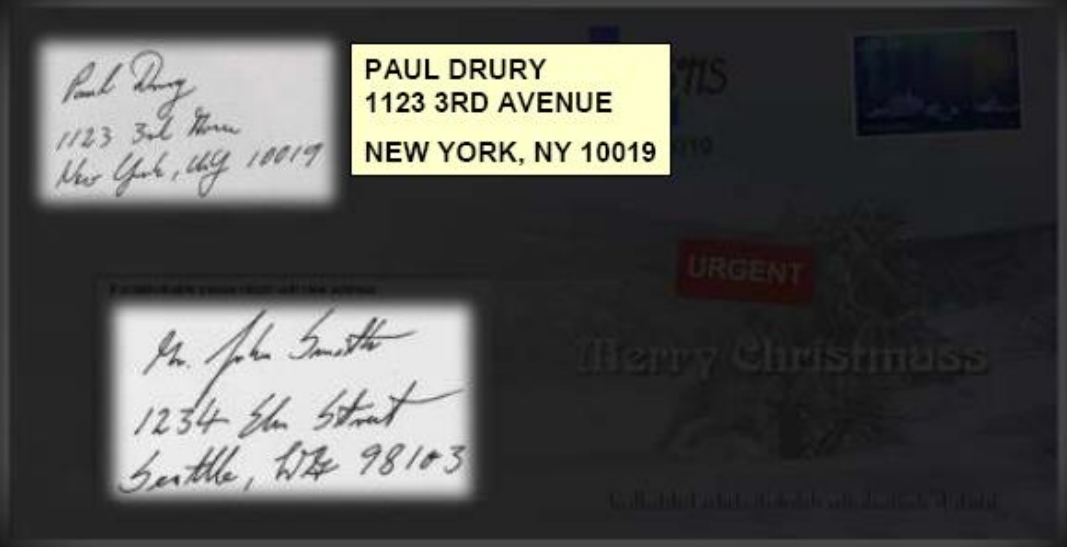
1982: First Address Reader Worldwide



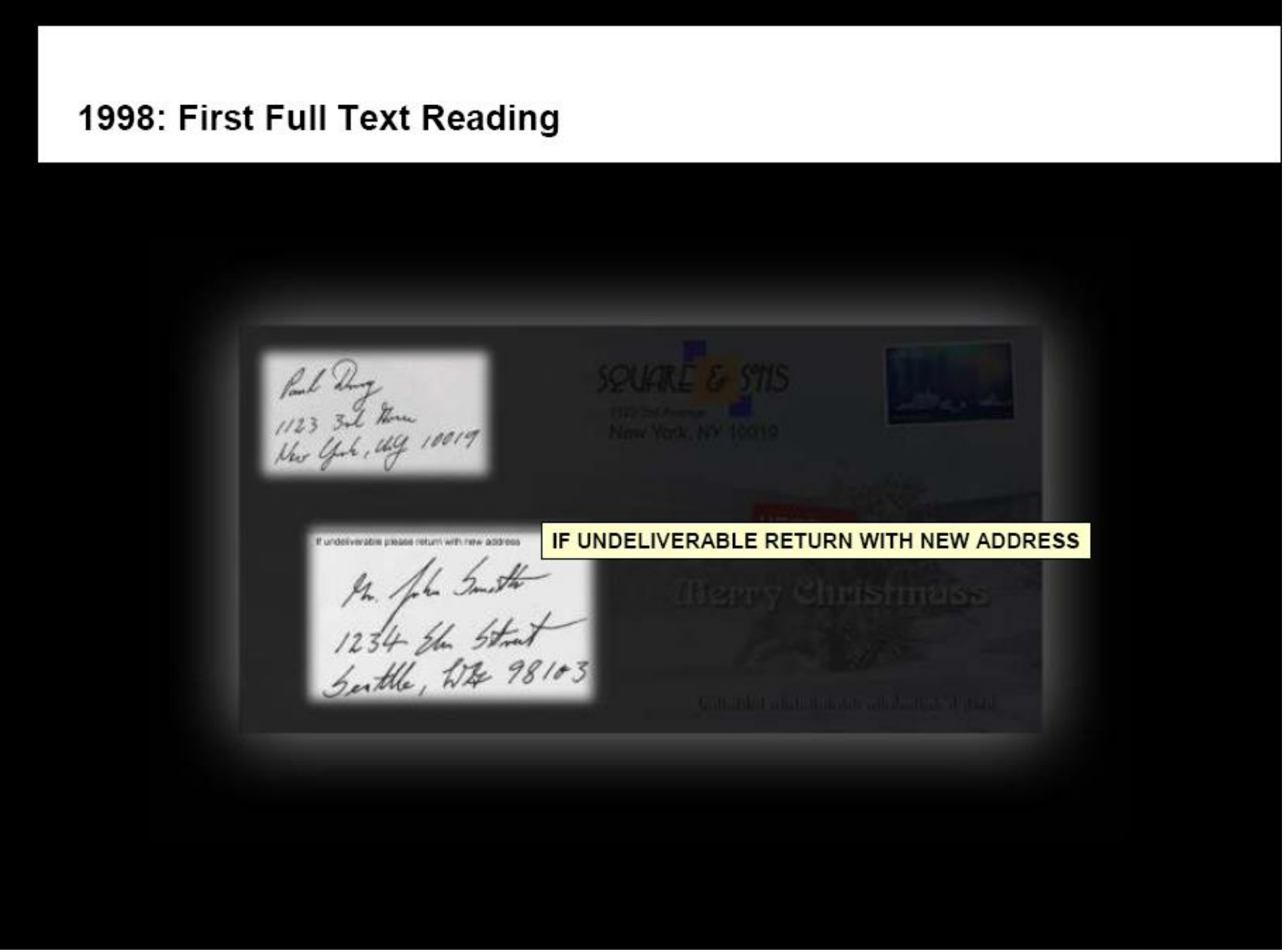
1984: First Multi Line Reader



1996: First Sender's Address Reader



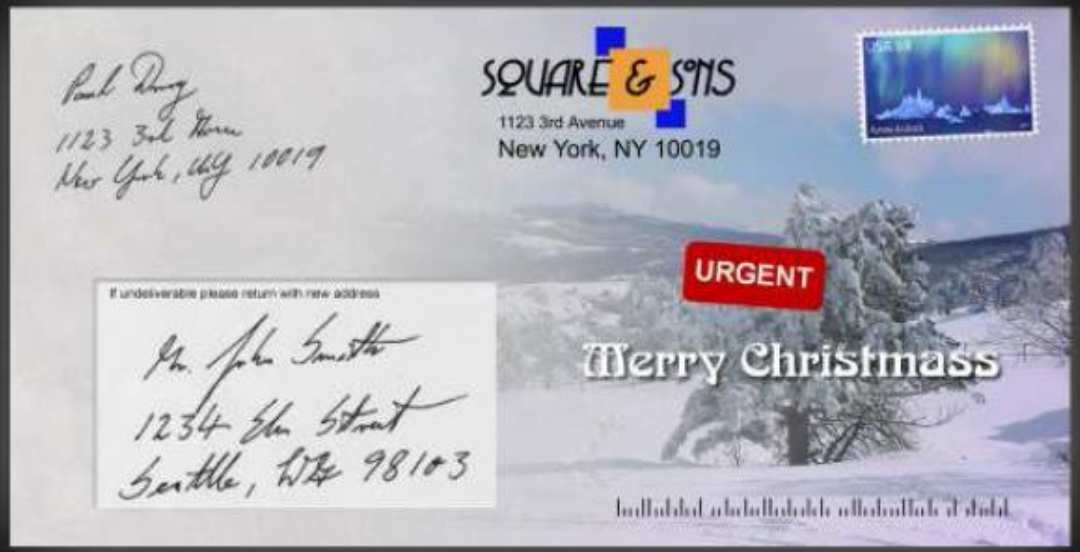
1998: First Full Text Reading



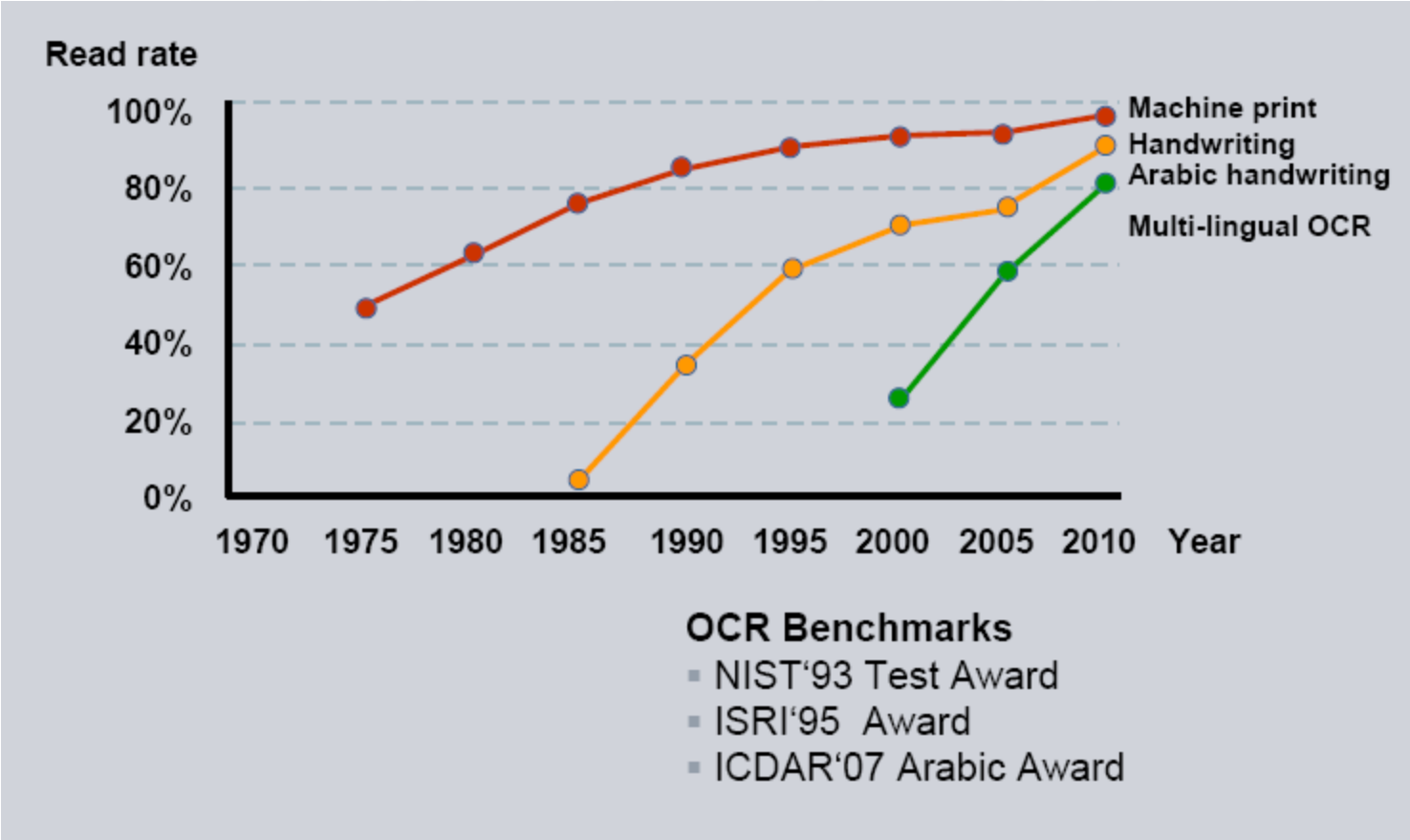
2000: First Graphics Recognition

2004: First Full Recognition

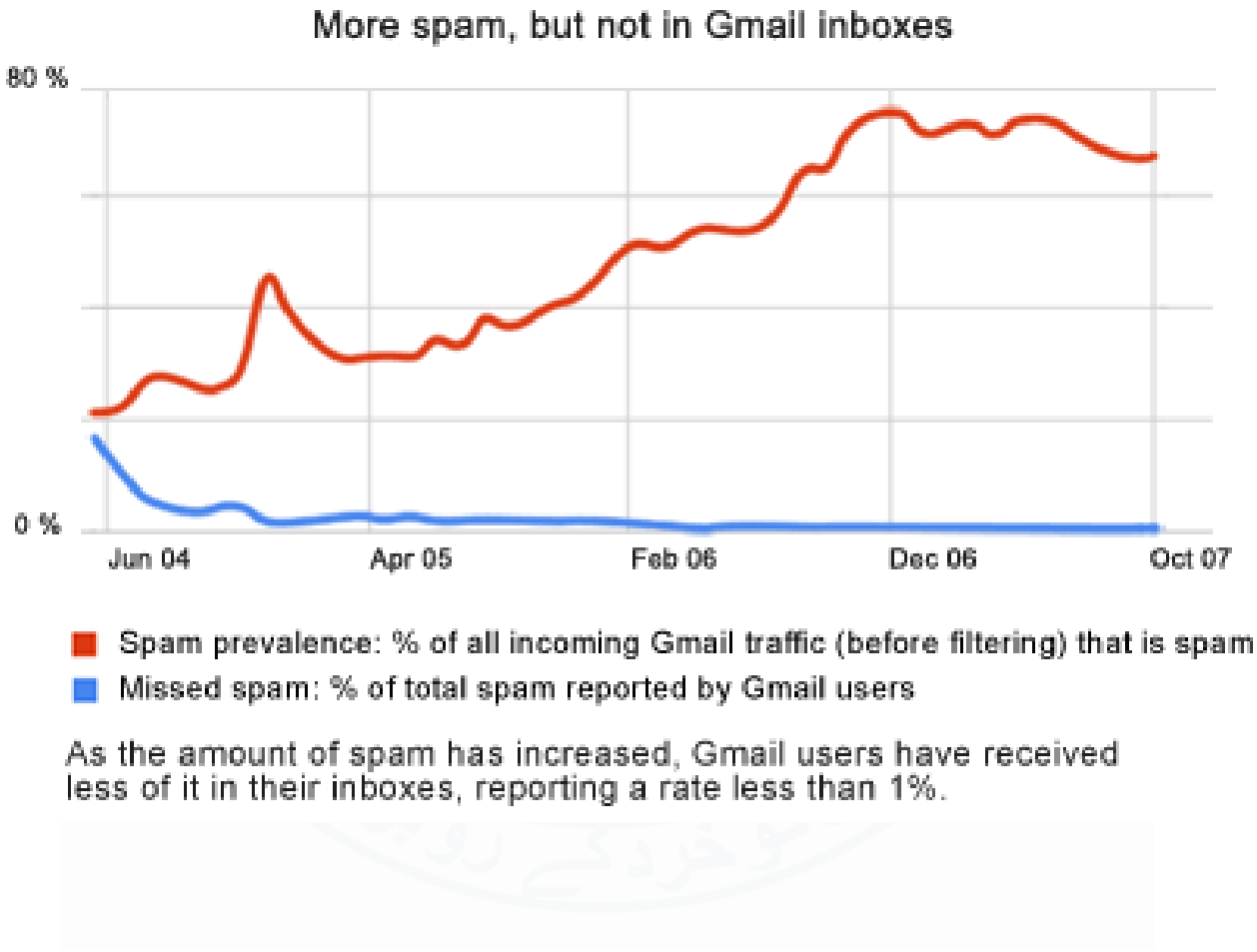
2008: Recognition on Both Sides of Envelope



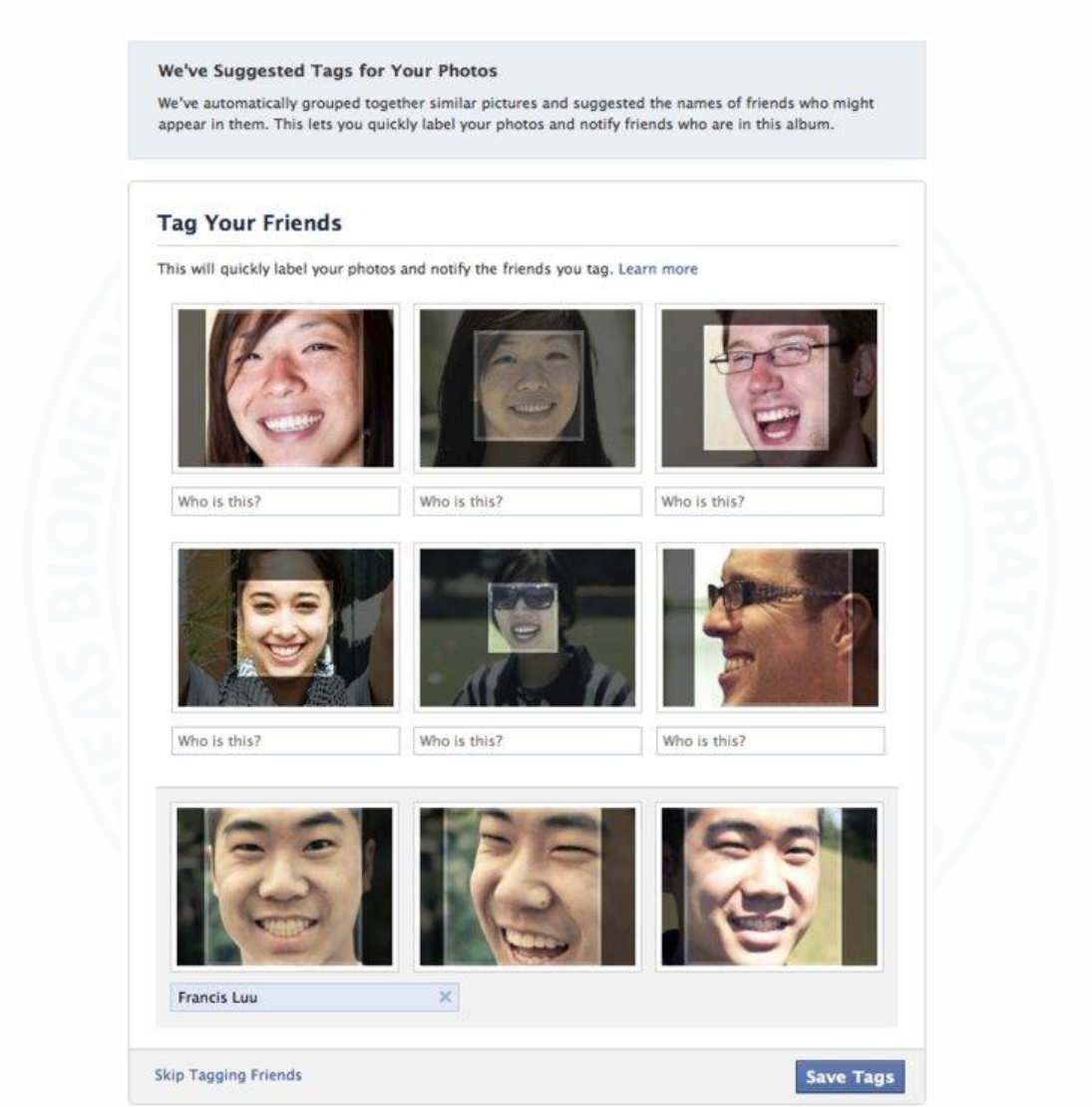
OCR Accuracy



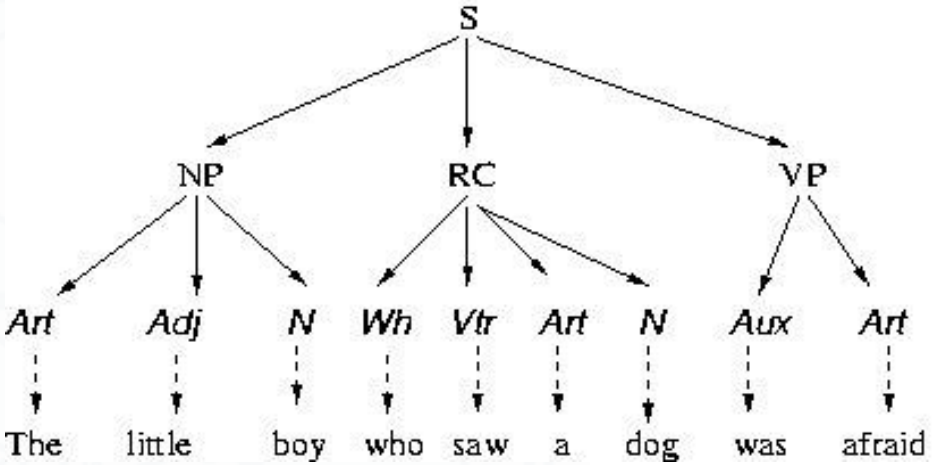
Gmail: ML in NLP

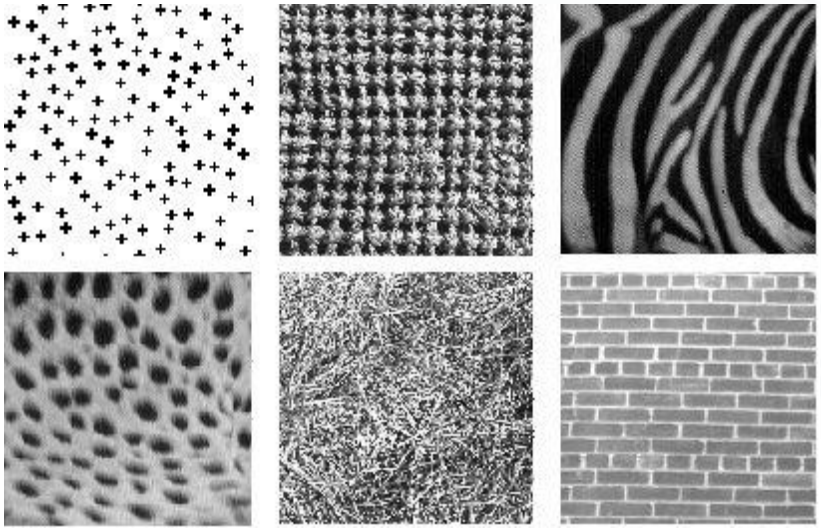


Facebook Friends Tagging



Applications of PR

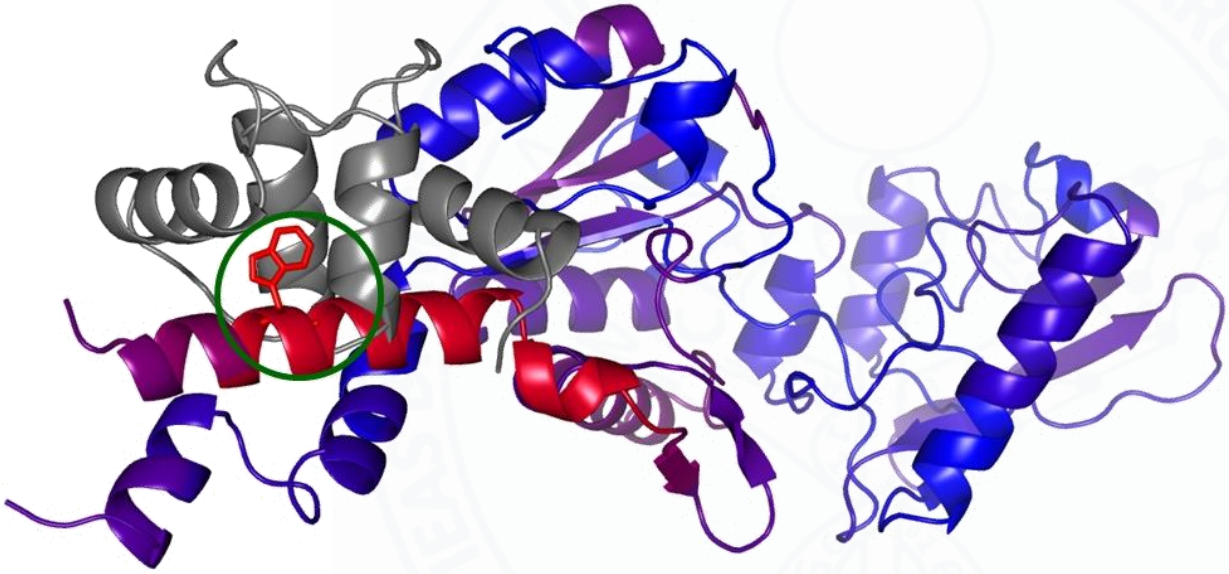


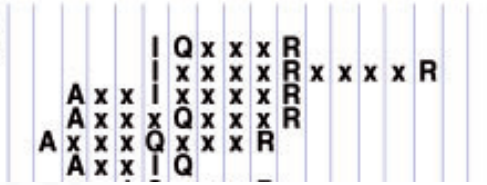


Recommender Systems

- Recommend movies based on user preferences, interests and likes
- Similar ideas for facebook…
  - Find friends that share your interests

Applications in Bioinformatics





High predicted binding propensity

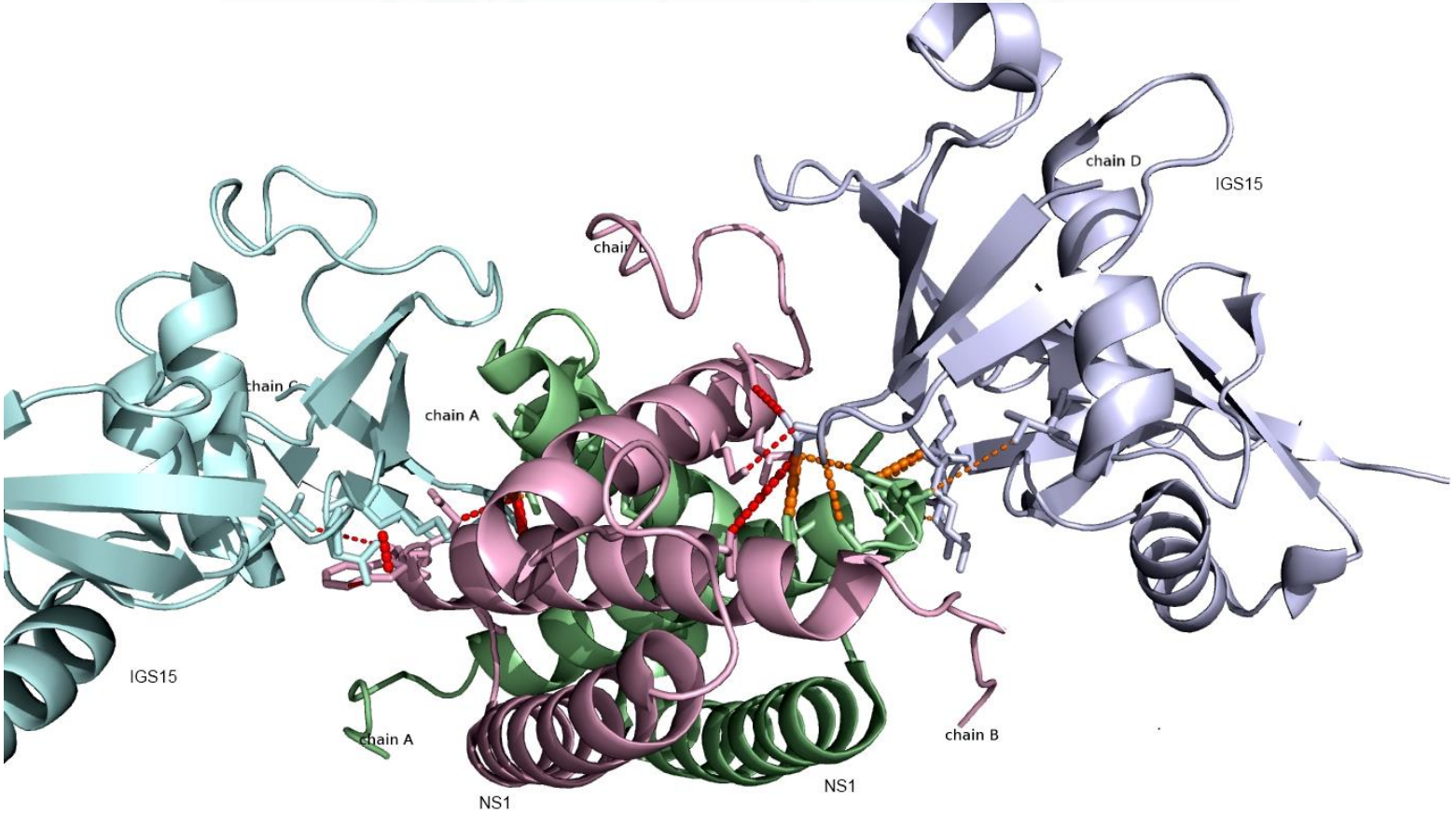
Low predicted binding propensity

Calmodulin

PDB: 1YRT

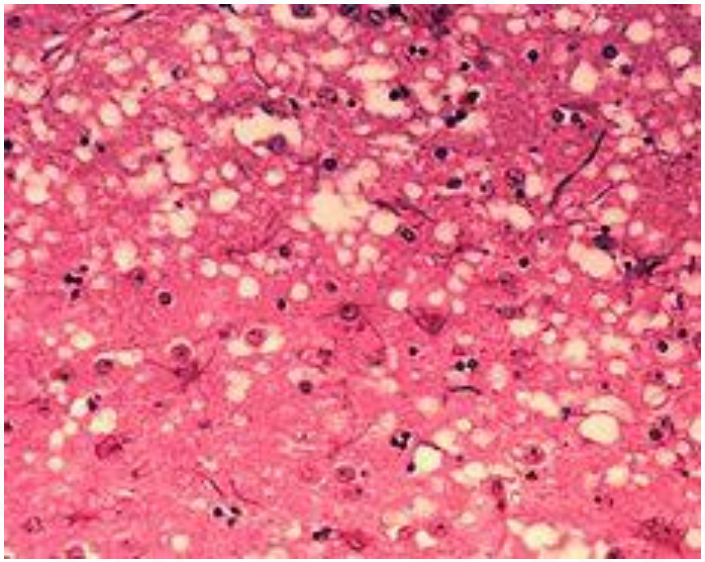
Applications in Bioinformatics

- Predict protein interfaces

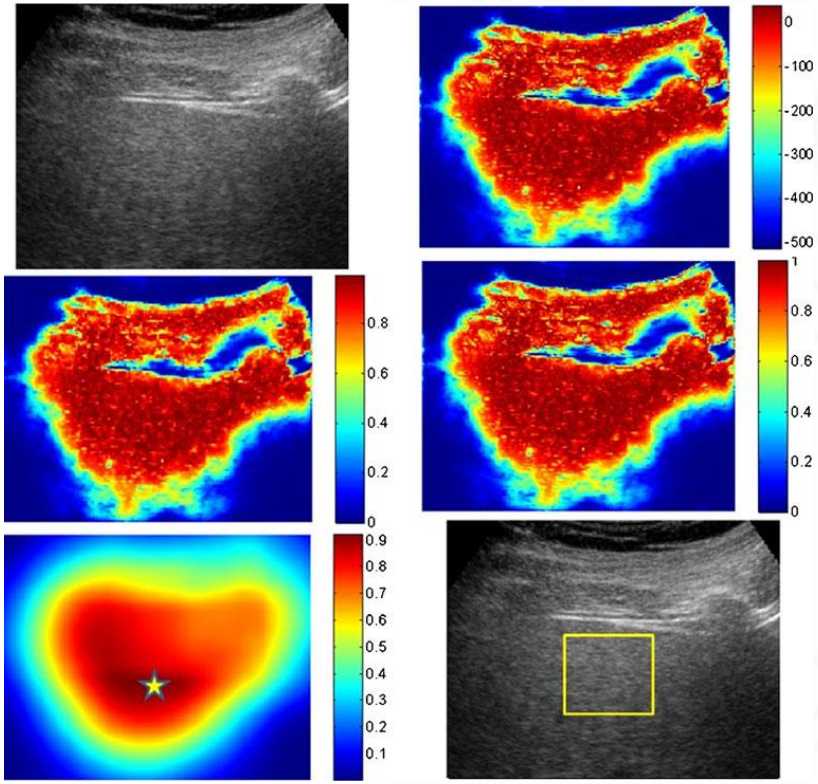


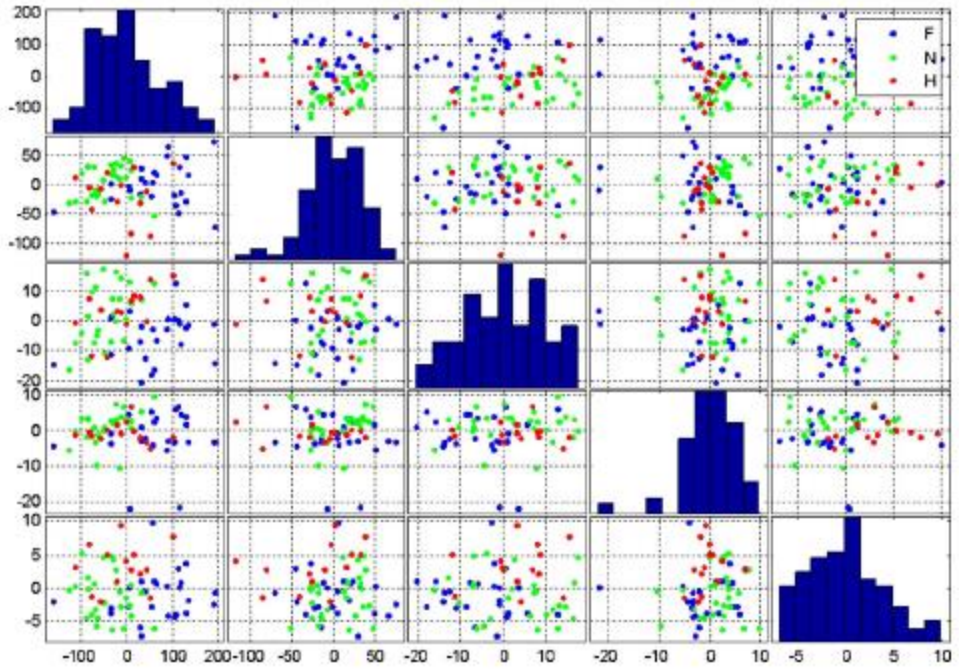
Applications in Bioinformatics

- Prion prediction using ML

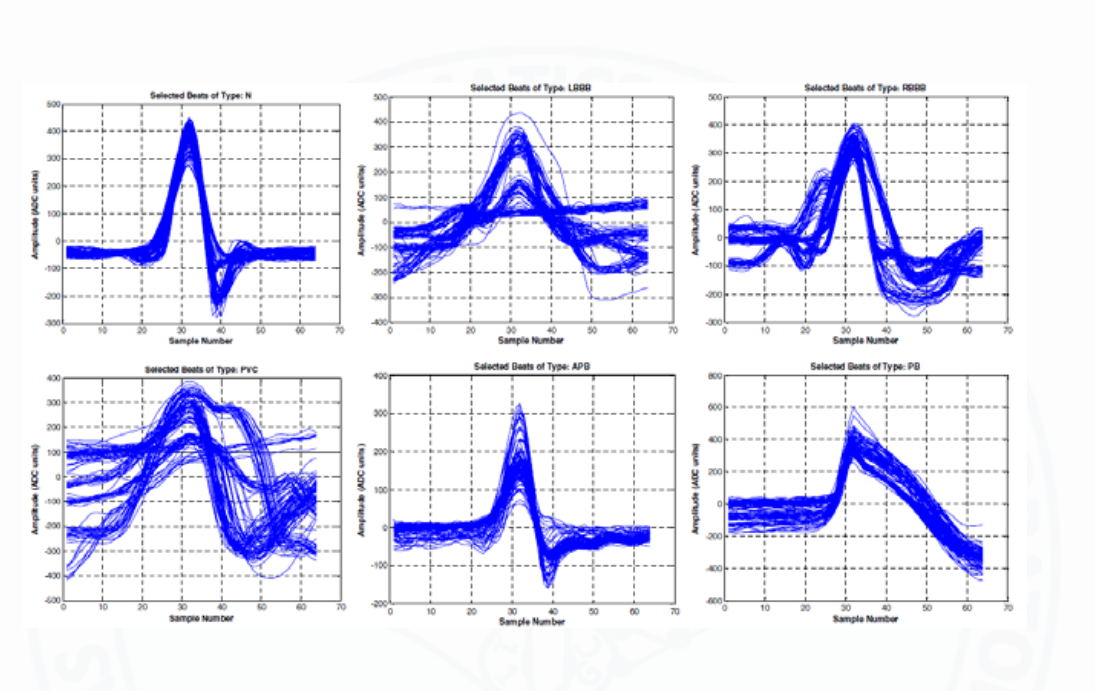


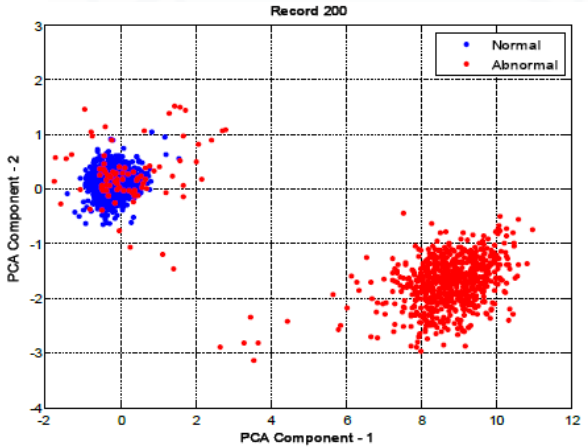
Medical Image Processing Applications

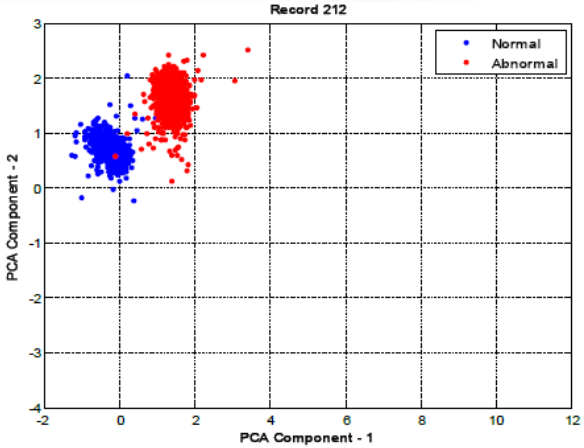




Applications in signal analysis







Computer / Network Security

- Prediction of threats
- Prediction of bugs / vulnerabilities in software
- Identification of malicious activity
- Identification of malicious software / viruses
- Attacking through side channels
  - Keyboard acoustics

Keyboard acoustics

Keyboard Acoustic Emanations Revisited

LI ZHUANG, FENG ZHOU, and J. D. TYGAR
University of California, Berkeley

Text recognized by the HMM classifier, with cepstrum features (underlined words are wrong),

> the big money fight has drawn the shoporo od dosens
> of companies in the entertainment industry as well
> as attorneys gnnerals on states, who fear the
> fild shading softwate will encourage illegal acyivitt,
> srem the grosth of small arrists and lead to lost cobs and
> dimished sales tas revenue.

Text after spelling correction using trigram decoding,

> the big money fight has drawn the support of dozens
> of companies in the entertainment industry as well as
> attorneys generals in states, who fear the film sharing
> software will encourage illegal activity, stem the growth
> of small artists and lead to lost jobs and finished sales
> tax revenue.

Original text. Notice that it actually contains two typographical errors, one of which is fixed by our spelling corrector.

> the big money fight has drawn the support of dozens
> of companies in the entertainment industry as well as
> attorneys gnnerals in states, who fear the file sharing
> software will encourage illegal activity, stem the growth
> of small artists and lead to lost jobs and dimished sales
> tax revenue.

PCR in HCI/CV

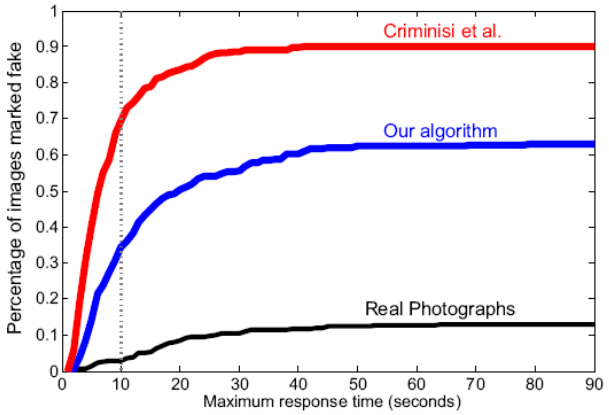
- Gesture Recognition

Scene Completion Using Millions of Photographs

James Hays Alexei A. Efros
Carnegie Mellon University

Original Image

Output

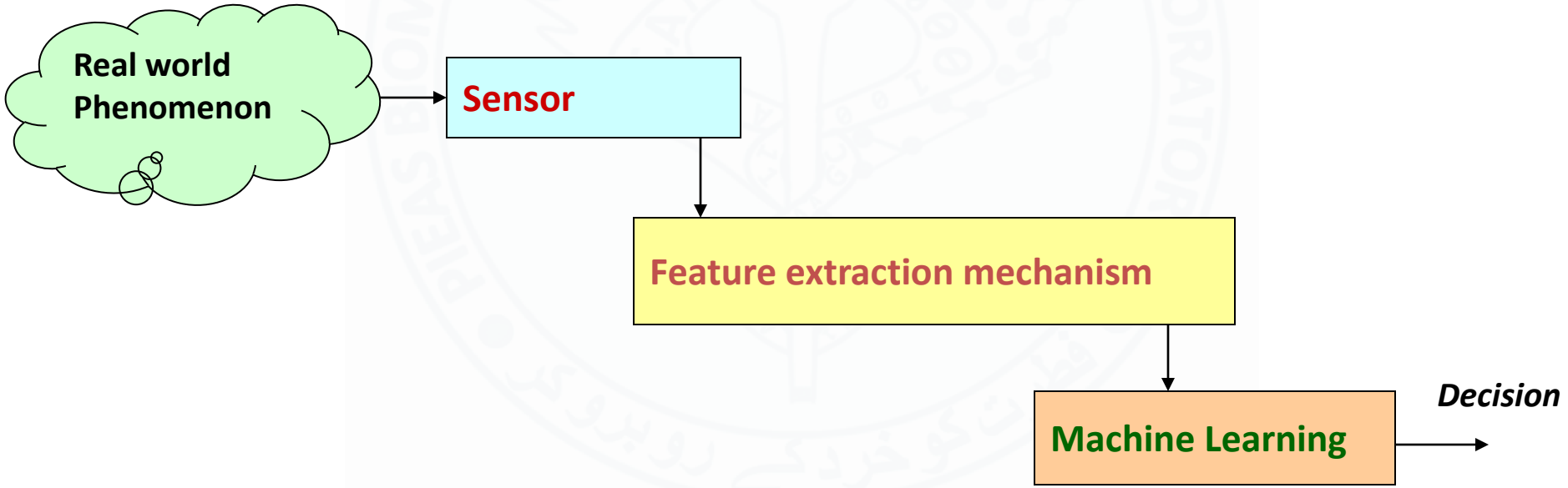


Deep Learning

- Google Brain
  - https://en.wikipedia.org/wiki/Google_Brain

Constructs of a PR System

- Identify the objective
  - Identify the unit of classification (example)
    - Image block, protein sequence, ….

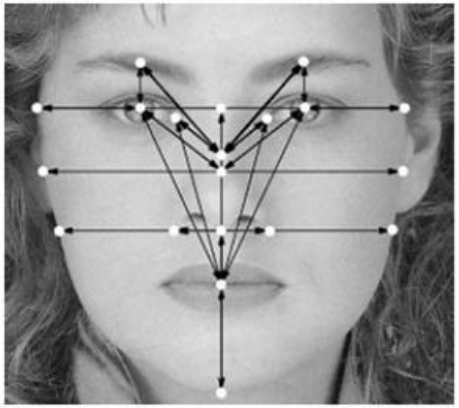


Components of a PR System

- Sensor
  - Responsible for getting raw data from an object
    - Examples
      - Camera for face recognition system
      - ECG for cardiac disease diagnosis
  - Multiple sensors can be combined to provide a better picture
  - Sensors can introduce noise into the PR system

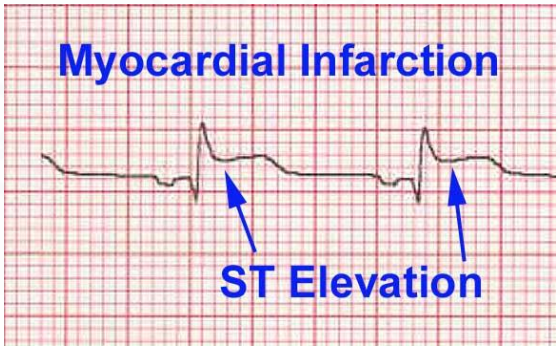
Components of a PR System…

- Feature Extraction
  - Usually (almost all the time!) the amount of raw data obtained from sensor(s) is too large and redundant
    - Example
    - 256x256 image acquired by a camera for face recognition contains 256x256x24 bits of data but what is the information we are looking for?
    - We are looking for features that purely characterize a face and as a result enable us to distinguish amongst faces



Components of a PR System…

- Feature Extraction…
  - Consider the design of a system for automated diagnosis of Myocardial Infarction
  - The usual sensor is an ECG machine which samples the electrical activity of the heart at some sample rate (e.g. 1000Hz)
  - We need some feature which would enable us to recognize particular ailments



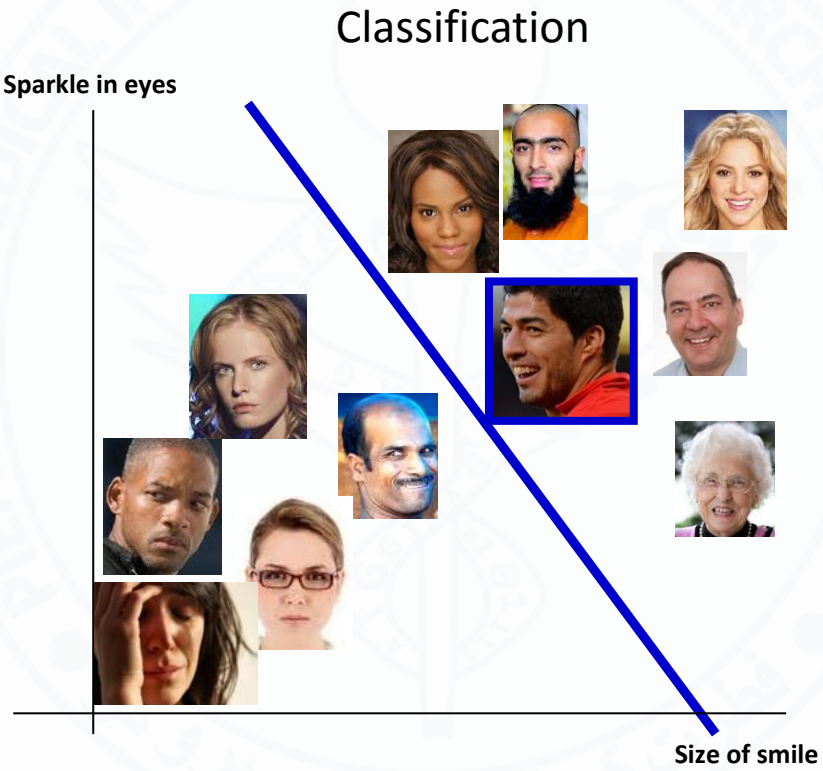
Components of a PR System…

- Feature Extraction…
  - Thus a feature extraction mechanism works as an information processor which takes in raw data and outputs information in the form of a feature vector which describes an object
  - We need descriptors called features which tend to remain somewhat constant over objects belonging to the same class but are different for objects belonging to other classes so we can discriminate easily between classes

Components of a PR System…

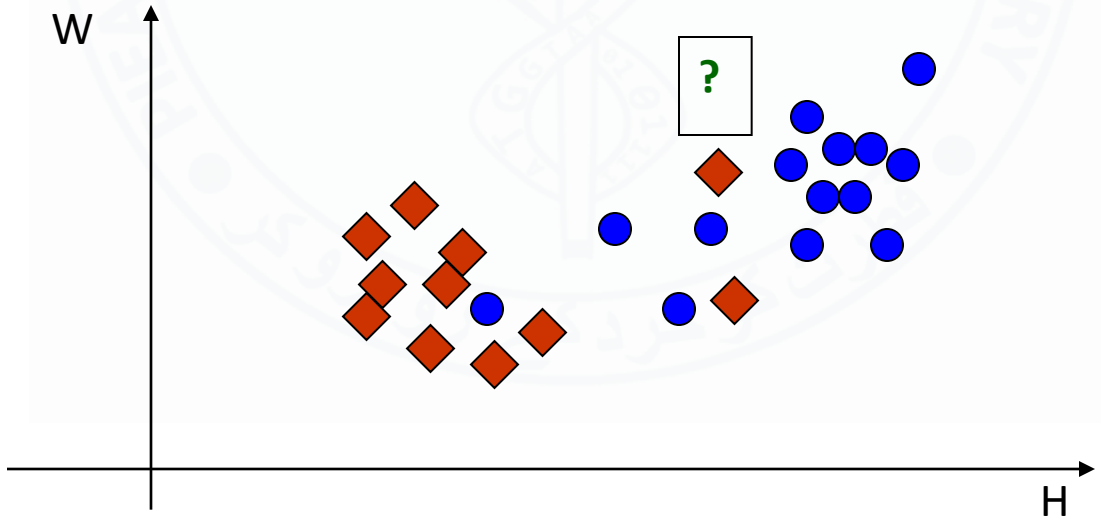
- Feature Extraction…
  - Points to note
    - A feature extraction mechanism
      - Extracts information in the form of a feature vector about an object which would enable our system to recognize it
      - Provides a form of dimensionality reduction aiming at removal of redundancy in data while maintaining discrimination between objects to be recognized
      - Computes numeric or symbolic information from the observations collected by sensor(s)
      - Helps the PR system to ignore noise effects of sensor(s)
      - Is problem specific

Example



Components of a PR System…

- Machine Learning
  - The job of the last stage of the PR system is to classify/describe objects on the basis of their features
  - Assigning labels to objects

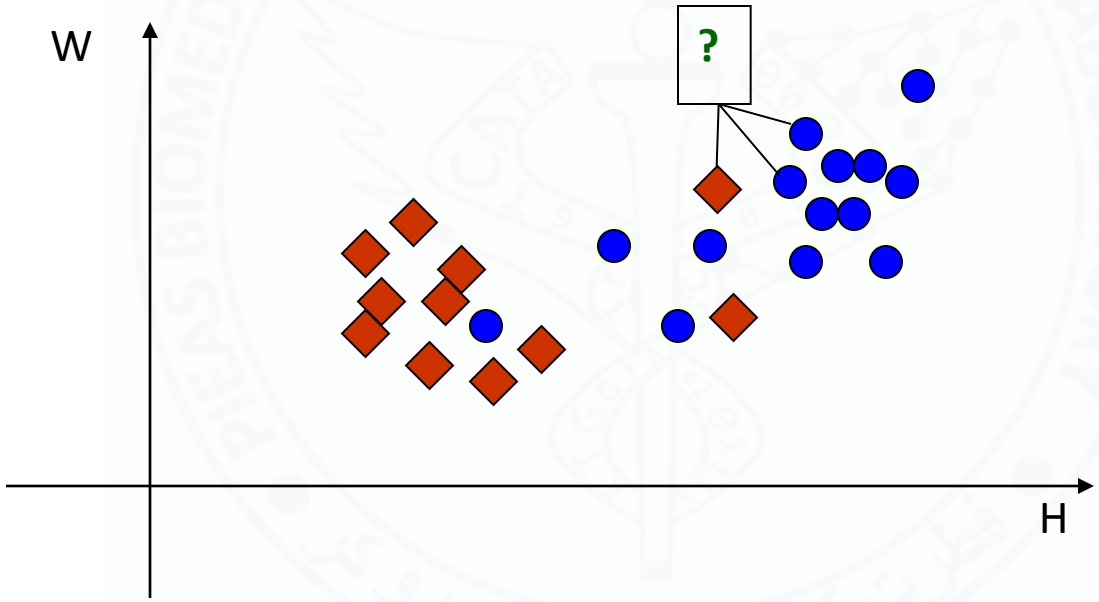


Components of a PR System…

- Classification/Description
  - Approaches
    - Using Apriori Knowledge
      - Use already know rules to make a decision
      - Example:
        - » ST deviation of more than 0.1mV in the ecg is indicative of Ischemia
    - Supervised Learning
      - Assumes that a set of already classified patterns (called the training set or training examples) is available and a learning strategy can be used to assign labels to unknown patterns (on the basis of the learning data)
        - » Input: A set of labeled examples (training feature vectors & their labels)
        - » Task: Find the boundary (discriminant) between classes
        - » Output: Given a unlabeled object, it uses the discriminant to assign a class label to it

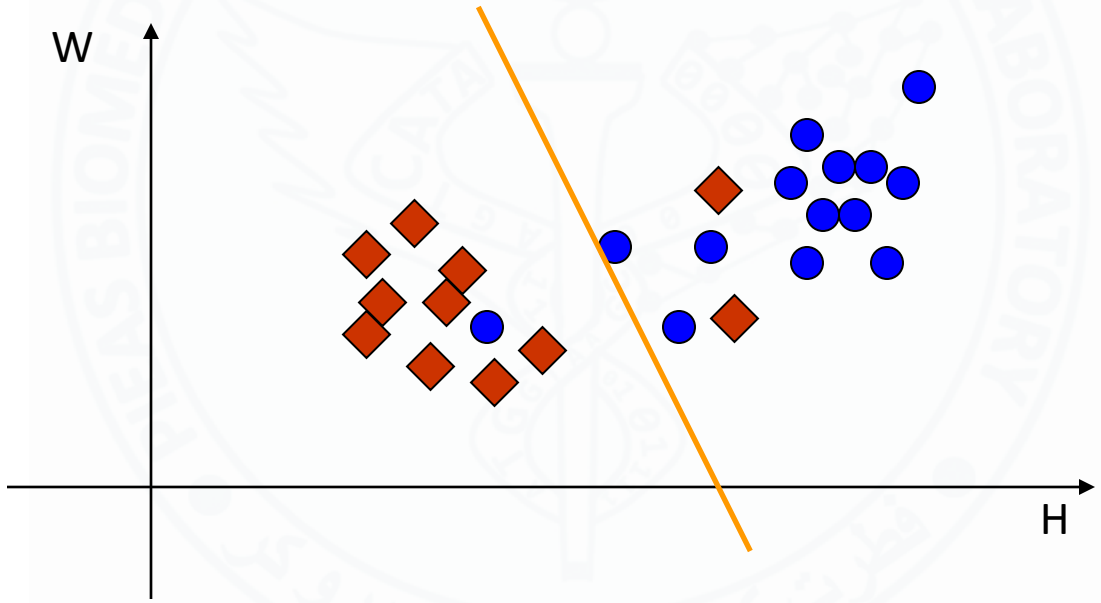
Classification Approaches: Supervised

- Example (k=3)-Nearest Neighbor Classification



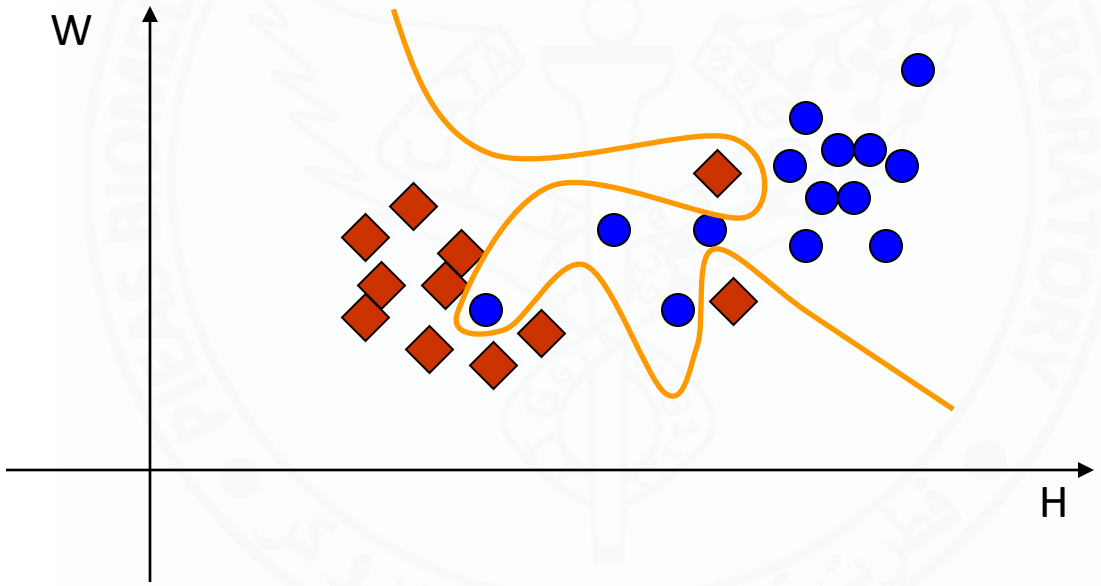
Classification Approaches: Supervised…

- Linear Classifier



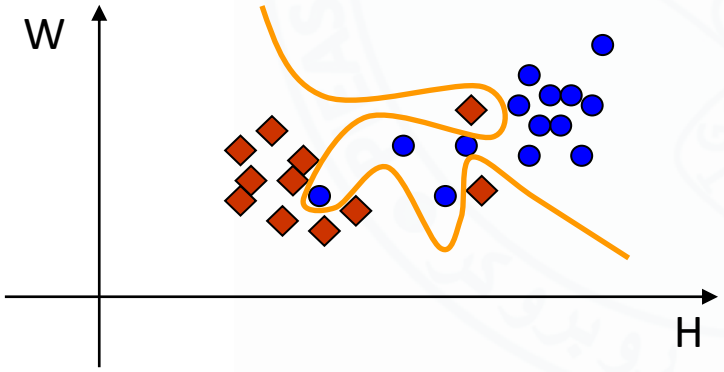
Classification Approaches: Supervised…

- Nonlinear Classification boundary

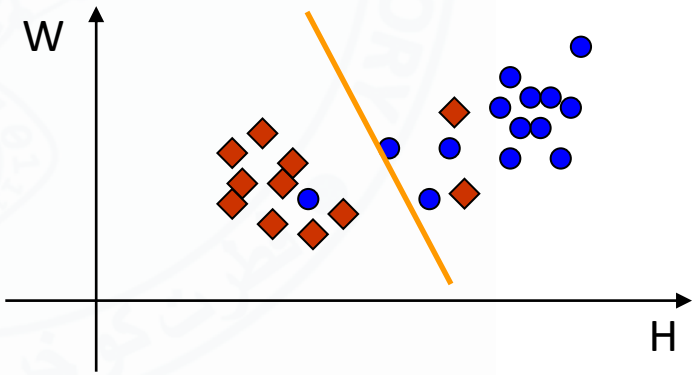


Classification Approaches: Supervised…

- Generalization vs. Memorization
  - A particular issue in classification is the tradeoff between memorization vs. generalization



Has great memorization but may generalize poorly



Has lesser memorization but may generalize better

Life Cycle


Collect data
Select features
Select model
Train classifier
Evaluate classifier

To Do

- Reading
  - Required
    - CHAPTER 1: Alpaydin, Ethem. *Introduction to Machine Learning*. Cambridge, Mass.: MIT Press, 2010.
    - Quiz Next Lecture
- Heads up!
  - Python Programming Assignment on Thursday

End of Lecture-2

We want to make a machine that will be proud of us.

- Danny Hillis